DIFFERENTIAL INCIDENCE OF TYPE II DIABETES MELLITUS BETWEEN MALE AND FEMALE USING BETA BINOMIAL MODEL

Matthew Aina [1], Muhammed Besiru Jibrin [2], Usman Abubakar Jauro [3], Olukayode Aiyeniko [4]

[1] Department of Statistics, Faculty of Physical Science, Ahmadu Bello University, Zaria, Nigeria
[2,3] Department of Computer Science, Federal University of Kashere, Gombe State, Nigeria
[4] Department of Computer Science, Faculty of Science, Lagos State University, Lagos State, Nigeria

Corresponding Author: Matthew Aina, matthew1969@yahoo.com

ABSTRACT: One of the life threaten diseases with a long duration and considerable associated morbidity is type II diabetes mellitus. This disease affects a large populations around the world regardless of economic status and its predominance is rising despite the struggles of clinicians, researchers and public health professionals. This paper intends to analysis gender sensitivity to type II diabetes mellitus. Secondary data used in this study was obtained for a period of 10 years from medical records of General Hospital, Ifaki-Ekiti, Ekiti state, Nigeria. A Beta-Binomial Model was applied by incorporating Bayesian Statistical Technique to determine the hypothesis testing, credible interval and posterior probability distribution of the incidence of type II diabetes mellitus between Male and Female patients. Experimental results revealed that male patients are more sensitive to type II diabetes mellitus, therefore, males are more affected with type II diabetes mellitus than female.

KEYWORDS: Beta-Binomial, Bayesian, Diabetes Mellitus, Morbidity, Sensitivity nature.

1. INTRODUCTION

Diabetes mellitus of type II can be defined as a long term metabolic disorder that is characterized by hyperglycemia (high blood sugar), disorders of carbohydrates, protein and fat metabolism ([B+13a]). The symptoms usually include increased thirst or hunger, recurrent urination and inexplicable loss of weight, feeling tired and sores that do not heal. The long-term complications from high blood sugar include heart disease, strokes, peripheral neuropathies, blindness, kidney failure, poor blood flow in the limbs which may lead to amputations. and renal disease ([CHF14]), The type-II diabetes mellitus has a very strong genetic component which is more greater than type I ([M+02]). There are no symptoms with is usually diagnosed by tests that indicate glucose intolerance and is treated with changes in diet and an exercise regimen.

The cases of diabetes worldwide have increased significantly in the last 3 decades ([B+13a]). It has been identified to be the fifth leading cause of death in the society nowadays ([W+04]). The prediction of studies revealed that in the next twelve years, type II diabetes mellitus will be the seventh leading cause of death in the society ([D+15]). In the developing nations, one in twenty deaths of adults is diabetes mellitus related [D+05]. It was observed that in the world, the region of African experience the incidence rate of diabetes mellitus at 3.8% a total of 13.1 million people [B+13b]. The problem of diabetes is increasing rapidly all over the globe causing an enormous socioeconomic and health challenge, researchers stated that the number of people with diabetes is expected to further increase from that of 415 million in 2015 to 642 million by 2040 [D+18].

There are several risk factors that lead to type II diabetes mellitus, these include unhealthy diet, physical inactivity, increasing age, high blood pressure, impaired glucose tolerance and poor nutrition during pregnancy. Type II diabetes mellitus makes up about 90% of cases of diabetes, with the other 10% due primarily to type I diabetes and gestational diabetes. Globally, there are major intervention sets to prevent incidence of type II diabetes mellitus. The interventions include dietary intervention or physical activity intervention or both. It has been discovered that combining dietary intervention and physical activity intervention is considered to be effective and efficient method for diabetes management.

Several studies have mentioned the impact of sexual role in diabetes mellitus in the general public, they have applied different approaches to carry out experimental and statistical analysis to investigate the effect of gender in chronic disease such as type II diabetes mellitus. This paper considered majorly the influence of gender on type II diabetes mellitus using beta binomial model to know the gender sensitivity to type 2 diabetes mellitus.

2. RELATED WORK

([D+18]) proposed a framework theoretically demonstrating the link between environment and type II diabetes mellitus. The study briefly

mentioned methodological challenges, potential solutions and future research chances. Walkability, air pollution, food, physical activity environment and roadways proximity were the most common environmental characteristics studied. Of the more than 200 reported and extracted relationships assessed in 60 studies, 82 showed significant association in the expected direction. In general, higher levels of walkability and green space were associated with lower type II diabetes mellitus, while increased levels of noise and air pollution were associated with greater risk. Current evidence is limited in terms of volume and study quality prohibiting causal inferences. However, the proof put forward that environmental characteristics may influence type II diabetes mellitus occurrence and also give a tangible basis for further investigation with better quality data and longitudinal studies with policy-relevant environmental measures.

[Nij17] applied negative binomial distribution to define the null hypothesis, to show the fraction of successes in each performed experiment and also the number of failure. The experiment showed that a good approximation for the negative binomial distribution is Poisson distribution, and it also established that the negative Binomial could be used to calculate the confidence interval percentage in based on the generalized hypergeometric function. The study recommended the areas of application of negative binomial such as health care, quality control, science and social science. It was further stressed that Negative Binomial may be applied in Aging research in clinical epidemiology, thunderstorm probability of occurrence in meteorology, stock control problems and accident statistics are good examples among other applications of negative Binomial distribution.

[ANH16] applied Bayesian Mixture Model Averaging (BMMA), which is a Bayesian solution to form the best single model from all the possible mixture models. BMMA. The modeling of BMM was used for Diabetes Mellitus data, while Bayesian Model Averaging was used for microarray data. The BMMA was applied to model blood sugar levels of Diabetes Mellitus patient through simulation studies. The experimental results of this study have succeeded in developing the BMMA normal model with two components mixture that could accommodate the real condition of the diabetes mellitus data with driven data concept. The model is capable of accommodating the real condition of Diabetes mellitus driven data concept in the future.

[B+15] employed Bayesian Logistic Regression Model (BLM) to model the risk factors associated with Diabetes Mellitus. The study compared the Bayesian logistic regression model with those from the Frequentist Logistic Regression Model (FLR) based on the significant factors. The results showed that BLM with non – informative flat prior distribution and Frequentist logistic regression model yielded similar results while non-informative flat model shared a different result compared to the model. The study suggested that the use of non-informative did not perfectly give improved model than the maximum likelihood estimate (MLE).

[JPJ14] applied negative regression to examine data on intravenous catheter (iv) insertion on sample of children. The study determined the fitness and applied negative binomial regression to examine data in southern stern U.S hospital. The assumptions governing negative binomial model was examined, that data may violate the assumption independence. The result showed that assumption of negative binominal is a good fit for modeling the (iv) insertion process data, with the AIC of the negative binomial regression model indicating an improved fit than the OLS regression model. The result revealed that negative binomial regression model appeared to give more precise effects.

[B+13a] presented a gender differential in the incidence of Diabetes mellitus among the patients at Udi village in Enugu State, Nigeria. The study formulated research questions and hypothesis, the method of descriptive survey design was applied. A sample of 343 diabetes mellitus patients were randomly selected and used for the study. Data documentary profoma was employed to collect the data from selected health facilities in respect to the patients' gender and their fasting sugar level as contained in their folder or hospital record/register. Frequency and percentages was used to answer research question while the null hypothesis was tested using Chi- square at 0.05 level of significance. The findings of the result showed a high incidence of diabetes mellitus in 2012, females recorded a high incidence of diabetes mellitus from 2008-2012 and high percentage of fasting blood sugar level than male. The findings also revealed no difference in the fasting blood sugar based on gender.

3. METHODOLOGY

In this proposed study, Beta Binomial model was used to carry out the analysis. This method shows how to update knowledge by applying Bayesian inference techniques. Beta binomial is a joint distribution which comprises beta as continuous distribution and Binomial as discrete distribution, which are mostly used in a Bayesian predictive model and can also be used in various aspects of inferential methods ranging from predictive, comparison and evaluation of experiments. The derivation of mean and variance of Beta Binomial joint distribution was shown, and some important terms related to Baye's theorem as shown in Equation (1).

$$h(\alpha \backslash x) = \frac{f(x \backslash \alpha) g(\alpha)}{\int f(x \backslash \alpha) g(\alpha)} \quad (1)$$

3.1 Data Collection

The study employed secondary data which was retrieved from patient's medical records for a period of 10 years at the General Hospital, Ifaki-Ekiti, Ekiti-State, Nigeria. From each patient's files the individual glucose level and record for both male and female patients were obtained. This helped to ascertain whether patients are pre-diabetic, if the blood glucose level rises to nearly 7.0mmo1/L and any patient whose blood glucose level rises above 7.0mmo1/L will be considered to be diabetic patients. The sample of the data used for this work is presented in Table 1.

Table 1: Sample of Data of Male and Female Patients affected with Type II Diabetes Mellitus (2007)

Gender / Month	Age					
	25-30		30 – 35		35 and above	
	M	F	M	F	M	F
January	2	3	3	2	4	3
February	3	1	1	-	2	2
March	-	1	2	3	3	1
April	1	1	1	1	4	-
May	-	2	2	1	3	2
June	-	2	3	2	5	1
July	-	2	1	-	1	1
August	-	-	4	-	4	-
September	1	2	2	3	4	-
October	4	1	2	2	-	3
November	1	-	1	1	4	-
December	2	-	3	-	5	2

3.2 Beta Binomial

Assume the prior distribution of P is Beta (α_1, α_2) and the conditional distribution of x given p is *Bin(n,p)* then,

$$f(x \ / \ p) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0,1, \dots, n \quad (2)$$

$$g(p) = \frac{|\overline{(\alpha_1 + \alpha_2)}}{|\overline{(\alpha_1)}|\overline{(\alpha_2)}} p^{\alpha_1 - 1} (1-p)^{\alpha_2 - 1}, \quad 0 < p < 1 \quad (3)$$

Therefore the compound distribution is

$$f(x \ / p)\, g(p) = \binom{n}{x} \frac{|\overline{(\alpha_1+\alpha_2)}}{|\overline{(\alpha_1)}|\overline{(\alpha_2)}} p^{x+\alpha_1-1} (1-p)^{n-x+\alpha_2-1} \quad (4)$$

The Bayesian approach will assist in extracting the features of interest from the posterior distribution. This work used Beta Binomial posterior distribution to determine the conditional probability variability between the genders susceptibility to type II diabetes mellitus. Bayesian Hypothesis test: The Bayesian compares hypothesis by comparing their posterior probability. All these hypothesis tests were done using STATA 14.2 version. Suppose the data x are *Bin(n,p)* and the prior is *Beta(α_1, α_2)*, so the posterior is *Beta($x+\alpha_1$; $n-x+\alpha_2$)*. The hypothesis can be stated as:

$$H_0 : p = \frac{1}{2}$$
$$H_1 : p > \frac{1}{2}$$

3.3 Credible interval

The credible intervals can be made by finding the marginal posterior distribution for the parameter of interest and find its $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quartile. *100(1-α)%* represents the interval between them. This study adopted the use of Beta Binomial distribution to estimate the interval between the genders susceptibility to type II diabetes mellitus. Another way to make credible interval is to find the marginal posterior distribution $h(\alpha \setminus x)$ for the parameter of interest and find the level set as discussed in Equation 5.

$$A_y = \{A \in \alpha : h(\alpha \setminus x) > y\} \quad (5)$$

4. RESULTS AND DISCUSSION

The phase presents overview of the analysis of this study and its findings. The results showed the differential analysis of incidence of type II diabetes mellitus between male and female by using beta binomial and negative binomial models to established which of these models is best and most efficient to fit differential incidence of type II Diabetes mellitus between Male and Female patients. The probability output of male and female diabetic patients is shown in Table 2 and Table 3.

Table 2: Probability Output of Male Diabetic Patients

Mean	0.01679
Std.DEV	0.01583
MCSE	0.00055
Median	-0.02405
95% Credible interval	0.000522-0.05723
BIC	140.308

Table 3: **Probability Output of Female Diabetic Patients**

Mean	0.01669
Std.DEV	0.01677
MCSE	0.00062
Median	- 0.01172
95% Credible interval	0.000381–0.061808
BIC	140.305

Tables 2 and 3 show that the posterior probability of male affected with type II diabetes mellitus to be 0.01679 while the posterior probability of female affected with type II diabetes mellitus to be 0.01669, which indicate that male have high chances of being affected with type II diabetes mellitus. Furthermore, Table 2 shows that the standard deviation of male affected with type II diabetes mellitus is 0.01583, while female affected with type II diabetes mellitus in Table 3, shows that the standard deviation is 0. 01677. .The standard deviation shows that males have low variability than female affected with type II diabetes mellitus. This indicates that male patients are more prone to type II diabetes mellitus than female patients. In addition, the median probability of male affected with type II diabetes mellitus (0.02405) is higher than the median of female affected with type II diabetes mellitus (0.01172). This shows that male patients have more chances of being susceptible to type II diabetes mellitus than female patients. Then Markov chain standard error (MCSE) for male patients shows lower variability of 0.00055 in comparison to female patient's variability of 0.00062. These results show that male patients have a great chance of being affected with type II diabetes mellitus than the female counterpart. The credible interval of 95% that male patients will be affected by type II diabetes mellitus is between 0.000522 to 0.05723 while female patients have 95% credible interval that falls between 0.000381 and 0.061808.

Finally, the Bayesian Information Criterion (BIC) that shows male patients will be affected with type II diabetes mellitus is 140.308 while female patients affected with type II diabetes mellitus, is 140. 305..This shows that males patients are likely to be affected with type II diabetes mellitus than female patients because BIC of male patients is higher than female patients BIC.

4.1 Trace Graph Representation of Type II Diabetes Mellitus (Females & Males)

The trace graph shows the probability of the male affected with type II diabetes mellitus to be 0.15 as discussed in Figure 1, while in the trace for probability that female is affected with type II diabetes mellitus to be 0.105 in the 9000 iteration is shown Figure 2. The study revealed that. Figure 1 the probability of trace of male probability was higher than trace female probability in Figure 2. It was observed that all the lines in Figure 2 fell below the grid line of probability of 0.105, which showed that females have less chances of being prone to type II diabetes mellitus compare to Figure 1 of male being affected with type II diabetes mellitus, because of the line that touched the grid line and this indicate that male patients have increasing chances of being affected with type II diabetes mellitus than the female patient's counterpart.

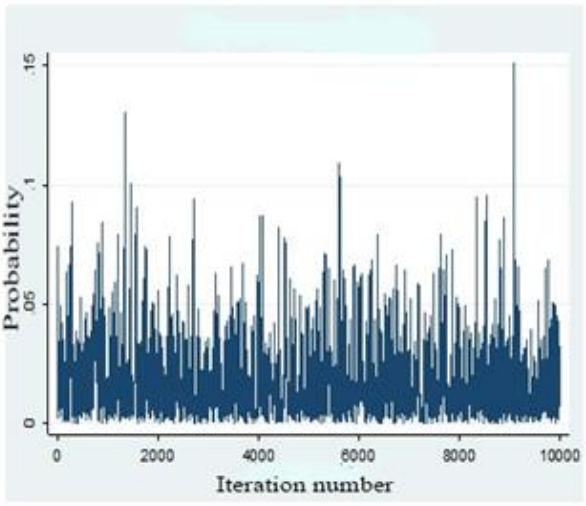

Figure 1: Trace of male diabetic patients

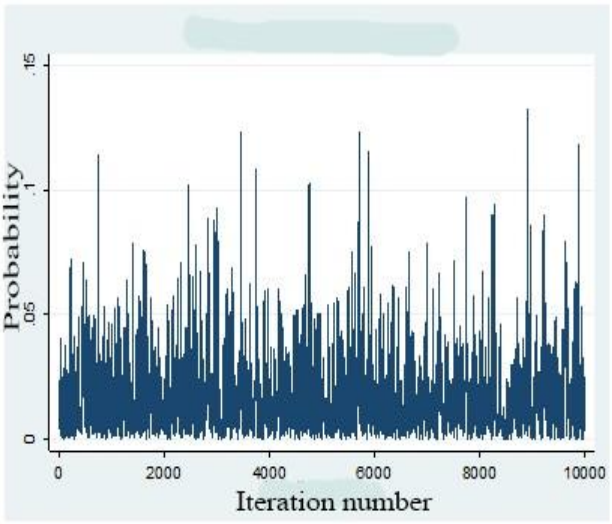

Figure 2: Trace of female diabetic patients

4.2 Histogram Presentation of Males and Females affected with Type II Diabetes Mellitus

This phase shows the graphical representation of comparison of histogram presentation of male and female patients affected with type II diabetes mellitus as shown in Figure 3 and Figure 4. The Histogram presentation indicated that male patients affected with type II diabetes mellitus is higher at

the zero probability by increasing above the grid lines and for female patients affected with type II diabetes mellitus the bar at the zero probability fell below the grid lines, this shows that male patients are more susceptible to type II diabetes mellitus than female patients.

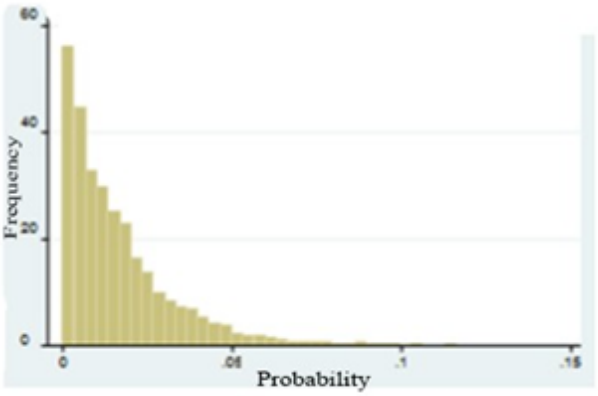

Figure 3. Histogram of Male Diabetic Patients

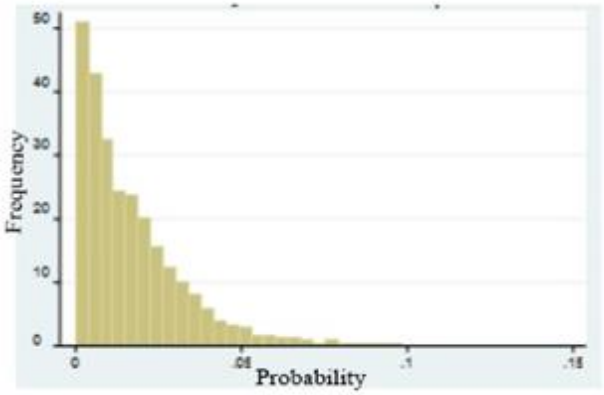

Figure 4. Histogram of Female Diabetic Patients

4.3 Density Graphical Presentation of Male and Female Patients Affected with Type II Diabetes Mellitus

The density graph for the posterior probability of male patients affected with the type II diabetes mellitus is 0.15 as shown in Figure 5 while the posterior probability of female affected with type II diabetes mellitus is 0.13 as shown in Figure 6, which is higher compared to female patients. The female patients have shown lower tendency of being affected with type II diabetes mellitus. The overall density line graph of Figure .5 and Figure 6 respectively starts from the frequency of 10 on the y-axis and they both slopes from left to right, but the line graph of Figure 5 stopped at probability of 0.15 for male patients and line graph of Figure 6 stopped at the probability of 0.13 for female patients. This result shows that male patients are more liable to type II diabetes mellitus than female patients respectively.

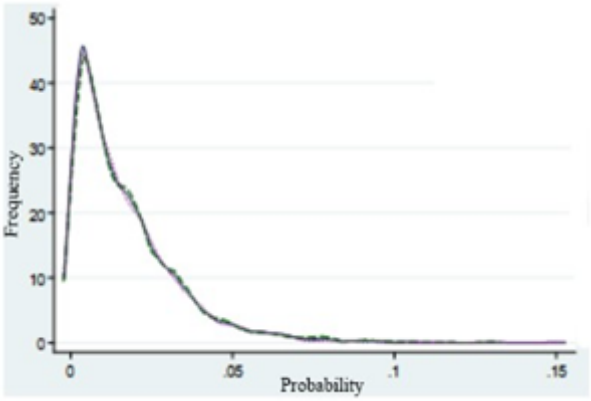

Figure 5. Density of male diabetic patients

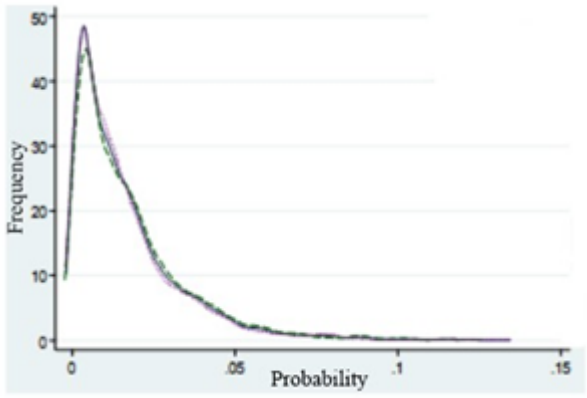

Figure 6. Density of female diabetic patients

CONCLUSION

Diabetes Mellitus has been identified as one of the life threaten diseases that occur in children and adults. The paper studied the influence of the gender sensitivity to type II diabetes mellitus. It was discovered that male patients are more prone to type II diabetes mellitus than female patients. The hypothesis testing of this study rejected the H_o, that male and female patients were equally affected by type II diabetes mellitus. The differential level of probability of female patients affected with type 2 diabetes mellitus which account for 0.135, at the 9000 iterations while male patients account for probability of 0.15 in the 9000 iteration. The Bayesian information criterion showed that male patients (140.308) is slightly higher than Bayesian information criterion of female patients (140.305), thus revealed that male patients more are liable to type II Diabetes Mellitus than female patients. Finally, based on Beta binomial joint distribution, which showed the comparison between male and female patients affected with type II Diabetes Mellitus, the study showed that male patients are more affected with type II Diabetes Mellitus than female patients.

REFERENCES

[ANH16] **K. L. Ani, B. A. Nur, I. Heri** - *Blood Sugar Levels of Diabetes Mellitus Patients Modeling with Bayesian Mixture Model Averaging*, Glob. J. Pure Appl. Math., vol. 12, no. 4, pp. 3158–3158, 2016.

[B+13a] **N. Bernice, O. Veronica, E. S. Samuel, C. Keyna** - *Gender Differential in the Incidence of Diabetes Mellitus among the Patients in Udi Local Government Area of Enugu State , Nigeria*, Mediterr. J. Soc. Sci., vol. 4, no. 8, pp. 131–138, 2013.

[B+13b] **S. Burner, E. Menchine, M. Taylor, E. Arora** - *Gender Differences in Diabetes Self-Management: A Mixed-Methods Analysis of a Mobile Health Intervention for Inner-City Latino Patients*, J. Diabetes Sci. Technol., vol. 7, no. 1, pp. 111–118, 2013.

[B+15] **B. Emenyonus, S. C. Mohd, B .A. Istrrinayagy, K. Shamarina, S. Chris** - *Bayesian Logistic Regression Model on Risk Factors of Type 2 Diabetes Mellitus*, Math. Theory Model., vol. 5, no. 1, pp. 2224–5804, 2015.

[CHF14] **M. Carolan, J. Holman, M. Ferrari** - *Experiences of diabetes self-management : a focus group study among Australians with type 2 diabetes*, J. Clin. Nurs., pp. 1–13, 2014.

[D+05] **I. Damnjanovic, D. Kitic, N. Stefanovic, S. Zlatkovic-Guberinic** - *Herbal self-medication use in patients with diabetes mellitus type 2*, Turkish J. Med. Sci., pp. 964–971, 2015.

[D+15] **H. Roglic, G. Unwin, N. Bennet, P. H. Mathers, C. Tumilehto, J. Nag, S. Connolly, V. King** - *The Burden of Mortality Attributable to*, Diabeties Care, vol. 28, no. 9, pp. 2130–2135, 2005.

[D+18] **T. Dendup, X. Feng, S. Clingan, T. Astell-Burt** - *Environmental Risk Factors for Developing Type 2 Diabetes Mellitus : A Systematic Review*, Int. J. Environ. Res. Public Heal., vol. 15, no. 78, pp. 1–25, 2018.

[JPJ14] **B. Jennifeer, M. Pamela, L. Jason** - *Exploring the use of Nagative Binomial Regression Modeling for Pediatric Peripheral Intravenous Catheterization*. J. Med. Stat. Inform., vol. 2, no. 6, pp. 2053–7662, 2014.

[M+02] **H. Meisinger, C. Thorand, B. Schneider, A. Stieber, J. Doring, A. Lowel** - *Sex Differences in Risk Factors for Incident Type 2 Diabetes Mellitus*, Arch Intern Med, vol. 162, pp. 82–89, 2002.

[Nij17] **V. Nijimbere** - *Hypothesis Test and Confidence Interval for the Negative Binomial Distribution Coincidence:A case For Rare Event*, International Journal of Contemporary Mathematical Science, 12 , 243-253. Retrieved from www.m_hikari.com, Int. J. Contemp. Math. Sci., vol. 12, pp. 243–253, 2017.

[W+04] **H. Wild. S. Green, A. Sicree, R. King** - *Global Prevalence of Diabetes: Estimate for the year 2000 and projections for 2030*, Diabtetes Care, vol. 27, pp. 1047–1053, 2004.